

SubstCRFsequencelisting_2009Jul30.txt

SEQUENCE LISTING

<110> KIM, Young Tae
 LEE, Jae Hyung

<120> Gene involved in the biosynthesis of carotenoid and marine
 microorganism, paracoccus haeundaesis, producing the
 carotenoid

<130> 428.1056

<140> US 10/551,508
 <141> 2006-06-05

<150> PCT/KR2004/000752
 <151> 2003-03-31

<150> KR2003-20222
 <151> 2003-03-31

<150> KR2003-20023
 <151> 2003-03-31

<160> 25

<170> KopatentIn 1.71

<210> 1
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> forward primer for Paracoccus haeundaesis 16S rDNA

<400> 1
 cataagtaat tatggttttg t 21

<210> 2
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> reverse primer for Paracoccus haeundaesis 16S rDNA

<400> 2
 cgcttcctta gaaaggag 18

<210> 3
 <211> 1454
 <212> DNA
 <213> Paracoccus haeundaesis

<400> 3
 caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa 60
 gtcgagcgag accttcgggt ctacgcggcg acgggtgagt aacgcgtggg aacgtgcctt 120

SubstCRFsequencelisting_2009Jul30.txt

```

tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tgggggaaag 180
atttatcgga gaaggatcgg cccgcgttgg attaggtagt tgggtgggta atggcccacc 240
aagccgacga tccatagctg gtttgagagg atgatcagcc aacttgggac tgagacacgg 300
cccagactcc tacgggaggg agcagtgggg aatcttagac aatgggggca accctgatct 360
agccatgccg cgtgagtgat gaaggcctta gggttgtaaa gctctttcag ctgggaagat 420
aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtataatcg 480
gaggggggcta gcgttggttcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag 540
tcagaggtga aatcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt 600
tcgagagagg tgagtggaat tccgagtgtg gaggtgaaat tcgtagatat tcggaggaac 660
accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcgaa agcgtgggga 720
gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc 780
aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg 840
caagattaaa actcaaagga attgacgggg gcccgcaaca gcggtggagc atgtggttta 900
attcgaagca acgcgcagaa ccttaccac ccttgacatg gcaggaccgc tggagagatt 960
cagctttctc gtaagagacc tgcacacagg tgctgcatgg ctgtcgtcag ctggtgtcgt 1020
gagatgttcg gttaagtccg gcaacgagcg caaccacgt ccctagttgc cagcattcag 1080
ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt 1140
tctcatggcc cttacggggtt gggctacaca cgtgctacaa tgggtggtgac agtgggttaa 1200
tccccaaaag ccatctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat 1260
cgctagtaat cgcggaacag catgccgagg tgaatacggt cccgggcctt gtacacaccg 1320
cccgtcacac catgggagtt ggttctaccc gacgacgctg cgctaacctt cggggggcag 1380
gcggccacgg taggatcagc gactgggggtg aagtcgtaac aaggtagccg taggggaacc 1440
tgcggtctga tcac 1454

```

```

<210> 4
<211> 6223
<212> DNA
<213> Paracoccus haeundaesis

```

```

<220>
<223> crt gene

```

```

<400> 4
gttccacgac tggggcatcc ccacgaccgc gtcgctgcgc gccatcgcgc cgatgatggg 60
gccggaccgg gttctggtcg ggtcgggcgg ggtgcgtcac gggctggacg ccgcgcgggc 120
catccgcctc ggcgcggaac tcgtggggca ggcggccgc gcgctgcccg ccgcgcgcca 180

```

SubstCRFsequencelisting_2009Jul30.txt

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| cagcgccgag | gccctgtccg | atcacctgtc | cgacgtcgtg | acccagctgc | gcatcgcgat | 240 |
| gttctgcacc | ggatcgggcg | accttgcagc | gctgcgctgc | gcgcctctgc | tggtgccggg | 300 |
| gccgggtggc | caatggtcgc | aagcaacggg | gatggaaacc | ggcgatgcgg | gactgtagtc | 360 |
| tgcgcggatc | gccggtcccg | gggacaagat | gagcgcacat | gccctgcca | aggcagatct | 420 |
| gaccgccacc | agcctgatcg | tctcgggcg | catcatcgcc | gcgtggctgg | ccctgcatgt | 480 |
| gcatgcgctg | tggtttctgg | acgcggcggc | gcatcccatc | ctggcgatcg | cgaatttcct | 540 |
| ggggctgacc | tggtctgcgg | tcggctctgtt | cttcatcgcg | catgacgcga | tgcacgggtc | 600 |
| ggtcgtgccg | gggcgtccgc | gcggcaatgc | ggcgatgggc | cagctgggcc | tgtggctgta | 660 |
| tgccggattt | tcgtggcgca | agatgatcgt | caagcacatg | gcccatacacc | gccataccgg | 720 |
| aaccgacgac | gaccccgatt | tcgacctagg | cgccccgggc | cgctgggtacg | cgcgcttcat | 780 |
| cggcacctat | ttcggctggc | gcgaggggct | gctgctgccc | gtcatcgtga | cggtctatgc | 840 |
| gctgatcctg | ggggatcgct | ggatgtacgt | ggtcttctgg | ccgtgccgt | cgatcctggc | 900 |
| gtcgatccag | ctgttcgtgt | tcggcacctg | gctgcgcac | cgccccggcc | acgacgcgtt | 960 |
| cccggaccgc | cataatgcgc | ggtcgtcgcg | gatcagcgac | cccgtgtcgc | tgctgacctg | 1020 |
| ctttcacttt | ggtggttatc | atcacgaaca | ccacctgcac | ccgacgggtgc | cttgggtggcg | 1080 |
| cctgcccgagc | acccgcacca | agggggacac | cgcatgacca | atttcctgat | cgtcgtcgcc | 1140 |
| accgtgctgg | tgatggagtt | gacggcctat | tccgtccacc | gttggatcat | gcacggcccc | 1200 |
| ctgggctggg | gctggcacia | gtcccaccac | gaggaacacg | accacgcgct | ggaaaagaac | 1260 |
| gacctgtacg | gcctggtctt | tgcggtgatc | gccacgggtgc | tgttcacggg | gggctggatc | 1320 |
| tgggcgccgg | tcctgtggtg | gatcgctttg | ggcatgaccg | tctatgggct | gatctatttc | 1380 |
| gtcctgcatg | acgggctggg | tcatcagcgc | tggccgttcc | gctatatccc | gcgcaagggc | 1440 |
| tatgcccgcc | gcctgtatca | ggccccaccg | ctgcaccacg | cggtcgaggg | acgcgaccat | 1500 |
| tgcgtcagct | tcggcttcat | ctatgcgcgg | ccggctcgaca | agctgaagca | ggacctgaag | 1560 |
| acgtcggggc | tgctgcgggc | cgaggcgag | gagcgcacgt | gacccatgac | gtgctgctgg | 1620 |
| caggggcggg | ccttgcgaac | gggctgatcg | ccctggcgct | gcgcgcggcg | cggccccgacc | 1680 |
| tgcgggtgct | gctgctggat | catgcggcgg | gaccgtcaga | cggccatacc | tggtcctgcc | 1740 |
| acgacccga | tctgtcgccg | cactggctgg | cgcggtgaa | gcccctgcgc | cgcgccaact | 1800 |
| ggccccgacca | ggagggtgcgc | tttccccgcc | atgccccggc | gctggccacc | ggttacgggt | 1860 |
| cgctggacgg | ggcggcgctg | gcggatgcgg | tggccccggc | gggcgcccag | atccgctgga | 1920 |
| acagcgacat | cggcctgctg | gatgaacagg | gggcgacgct | gtcctgcggc | acccggatcg | 1980 |
| aggcgggcgc | ggtcctggac | gggcgcggcg | cgcagccgtc | gcggcatctg | accgtgggtt | 2040 |
| tccagaaatt | cgtgggcgtc | gagatcgaga | ccgactgccc | ccacggcggtg | ccccgcccga | 2100 |

SubstCRFsequencelisting_2009Jul30.txt

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|------------|------|
| tgatcatgga | cgcgaccgtc | accagcagg | acgggtaccg | attcatctat | ctgctgccct | 2160 |
| tctctccgac | gcgcatcctg | atcgaggaca | ctcgctattc | cgatggcggc | aatctggacg | 2220 |
| acgacgcgt | ggcggcggcg | tcccacgact | atgcccgcga | gcagggctgg | accggggccg | 2280 |
| agggtccggcg | cgaacgcggc | atcctgcccga | ttgcgctggc | ccatgacgcg | gcgggcttct | 2340 |
| gggccgatca | cgcgaggagg | cctgttcccc | tgggactgcg | cgcggggttc | tttcaccccg | 2400 |
| tcaccggcta | ttcgctgccc | tatgcggcgc | agggtggcga | cgtggtggcg | ggcctgtccg | 2460 |
| ggccgcccgg | caccgacgcg | ctgcgcggcg | ccatccgcga | ttacgcgatc | gaccgggcac | 2520 |
| gccgtgaccg | ctttctgcgc | ctgctgaacc | ggatgctggt | ccgcggctgc | gcgcccgacc | 2580 |
| ggcgtatac | cctgctgcag | cggttctacc | gcatgccgca | tggactgatc | gaacggttct | 2640 |
| atgccggccg | gctgagcgtg | gcggatcagc | tgcgcatcgt | gaccggcaag | cctcccatte | 2700 |
| cccttggcac | ggccatccgc | tgccctgccc | aacgtcccct | gctgaaggaa | aacgcatgaa | 2760 |
| cgcccattcg | cccgcggcca | agaccgccat | cgtgatcggc | gcaggctttg | gcgggctggc | 2820 |
| cctggccatc | cgccctgcag | ccgcggggcat | cgccaccacc | ctggctcagg | cccgggacaa | 2880 |
| gcccggcggg | cgcgccatg | tctggcacga | tcagggccat | gtcttcgacg | cgggcccgcg | 2940 |
| cgtcatcacc | gaccccgatg | cgctcaagga | gctgtggggc | ctgaccgggc | aggacatggc | 3000 |
| gcgcgacgtg | acgctgatgc | cggtgtcgcc | cttctatcga | ctgatgtggc | cgggcgggaa | 3060 |
| ggtcttcgat | tacgtgaacg | aggccgatca | gctggagcgc | cagatcggcc | agttcaaccc | 3120 |
| ggacgacctg | gaaggatacc | gccgcttccg | tgattacgcg | gaggagggtg | atcaggaggg | 3180 |
| ctacgtcaag | ctgggcaccg | tgcccttcct | caagctgggc | cagatgctca | aggccgcgcc | 3240 |
| cgcgctgatg | aagctggagg | cctataagtc | cgtccatgcc | aaggctcgca | ccttcatcaa | 3300 |
| ggacccctat | ctgcggcagg | cgttttcgta | tcacacgctg | ctggtggggc | ggaatccctt | 3360 |
| ctcgaccagc | tcgatctatg | cgctgatcca | cgcgctggag | cggcgcggcg | gggtctggtt | 3420 |
| cgccaagggc | ggcaccaacc | agctggtcgc | gggcatggtc | gcgctgttcg | aacggcttgg | 3480 |
| cggccagatg | atgctgaacg | ccaaggctgc | ccggatcgag | accgagggcg | cgcggaccac | 3540 |
| gggcgtcacc | ctggcggacg | ggcggctctt | aagggccgac | atggtcgcca | gcaacggcga | 3600 |
| cgtcatgcac | aactatcgcg | acctgctggg | ccacacggcc | cgcgggcaga | gccgcgcgaa | 3660 |
| atcgctggac | cgcaagcgct | ggtccatgtc | gttggttcgtg | ctgcatttcg | gtctgcgcga | 3720 |
| ggcgcccaag | gacatcgcg | atcacaccat | cctgttcggc | ccccgctaca | gggagctggt | 3780 |
| caacgagatc | ttcaagggcc | cgaagctggc | cgaggatttc | tcgctgtacc | tgcattcgcc | 3840 |
| ctgcacgacc | gatccggaca | tggcgctcc | gggcatgtcc | acgcattacg | tgctggcccc | 3900 |
| cgtgccgcat | ctgggcccgc | ccgagatcga | ttgggcggtc | gaggggcccgc | gctatgccga | 3960 |

SubstCRFsequencelisting_2009Jul30.txt

ccgcatcctg gcgctccctgg aggagcggct gatcccgaac ctgcgcgcca acctgaccac 4020
 gacgcgcatac ttcacgcccc cggatttcgc cagcgaactg aacgcccatac acggcagcgc 4080
 cttctcgggtc gagccgatcc tgacgcaatc cgcgtgggttc cggccgcaca accgcgacaa 4140
 gacgatccgc aactttctatc tggtcggcgc gggcacccat ccgggcgcgg gcattccggg 4200
 cgtcgtgggc tcggccaagg ccacggccca ggtgatgctg tccgacctgg cgggcgcatac 4260
 agcgatctgg tcctgacctc gaccgaggcg atcacccaag ggtcgcaaag ctttgccacg 4320
 gcggccaagc tgatgccgcc gggcatccgc gacgacacgg tgatgctcta tgcctggtgc 4380
 cgccacgcgg atgacgtgat cgacggtcag gccctgggca gccgccccga ggcggtgaac 4440
 gacccgcagg cgcggctgga cggcctgcgc gtcgacacgc tggcggccct gcagggcgac 4500
 ggtccggtga ccccgccctt tgccgcgctg cgcgcggtgg cgcggcggca tgatttcccc 4560
 caggcctggc ccatggacct gatcgaaggc ttcgcgatgg atgtcgaggc gcgcgactat 4620
 cgcacgctgg atgacgtgct ggaatatctc tatcacgtcg caggcatcgt cggcgtgatg 4680
 atggcccgcg tgatgggcgt gcgcgacgat cctgtcctgg accgcgcctg cgacctgggg 4740
 ctggcgttcc agctgaccaa catcgcgcgc gacgtgatcg acgatgcgcg catcgggcgg 4800
 tgctatctgc cgggggactg gctggaccag gcgggcgcgc ggatcgacgg gccggtgccg 4860
 tcgccggagc tgtacacagt gatcctccgg ctgttgatg aggcggaacc ctattacgcg 4920
 tcggcgcggg tgggtctggc ggatctgcca ccgcgctgcg cctggtccat cgccgccgcg 4980
 ctacggatct atcgcgccat cgggctgcgc atccgcaaga gcgggccgca ggcctatcgc 5040
 cagcggatca gcacgtccaa ggctgccaag atcggcctgc tgggcgtcgg gggctgggat 5100
 gtcgcgcgat cacgcctgcc gggggcgggc gtgtcgcggc agggcctctg gacccggccg 5160
 catcacgtct aggcgcgcgc ggcgtagggc agaaccggtt ccagcagggc cgcgatttcc 5220
 ggagcctgaa ggcgcttgct gcgcagcatc gcgtccagtt gggcgcggtt ggcctcgtaa 5280
 tgacgggaca cgttctgcag gtctgacacg gccagaaggc cgcgccgcgg gccggggggc 5340
 gcggcatcgc gaccggtatc cttgccaagc gccgcctggt cggccacgac gtccagcagg 5400
 tcgtcatagg actggaacac gcggcccagc tgacggccaa agtcgatcat ctgggtctgc 5460
 tcctcggcgt cgaactcctt gatcacggcc agcatctcca gcccgcgat gaacagcacg 5520
 ccggtcttca ggtcctgttc ctgttcgacc cccgcgccgt tcttggccgc gtgcagggtc 5580
 aggtcctggc cggcgcacag gccctgcggc cccagggacc gcgacaggat ccgcaccagc 5640
 tgcgcccga cgtgcccga cgcgccgcgc gcaccggcca gcagggccat tgcctcggtg 5700
 atcagggcga tgccgccag cacggcacgg ctttcgccat gcgccacatg ggtcgcgggc 5760
 cggccgcggc gcagccggc atcgtccatg cagggcaggc cgtcgaagat cagcgatgcg 5820
 gcatgcacca tctcgaccgc gcaggcggcg tcgacgatcg tgtcgcagac cccgcccag 5880

SubstCRFsequenceListing_2009Jul30.txt

```
gcctctgccg caagcagcat cagcatgccg cggaaccgcc tgcccgcga cagcgcgcca 5940
tggctcatgg ccgcgccgag cggctgcgac acggcaccga atccctgggc gatctcctca 6000
agtctggtct gcagaagggt ggcgtggatc gggttgacgt ctcgtctcat cagtgccttc 6060
gcgcttgggt tctgacctgg cgggaagggtc aggccggggc ggcaccccgt gaccggtcat 6120
ccaccgtcaa cagtccccat gttggaacgg ttcacgcccg attgcgagcc ttttcgacgg 6180
cgacgcgggg tcgcgcggca atttgtccaa caaggtcagt gga 6223
```

```
<210> 5
<211> 729
<212> DNA
<213> Paracoccus haeundaesis
```

```
<220>
<223> crtW gene
```

```
<400> 5
atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc 60
ggcatcatcg ccgcgtggct ggccttgcac gtgcatgcgc tgtggtttct ggacgcggcg 120
gcgcattcca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtcgggtctg 180
ttcttcatcg cgcattgacg gatgcacggg tcgggtcgtg cggggcggtc gcgcggcaat 240
gcggcgatgg gccagctggt cctgtggctg tatgccggat tttcgtggcg caagatgatc 300
gtcaagcaca tggcccatca ccgccatacc ggaaccgacg acgaccccga tttcgaccat 360
ggcggcccg tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg 420
ctgctgctgc ccgtcatcgt gacggtctat gcgctgatcc tgggggatcg ctggatgtac 480
gtggtcttct ggccgctgcc gtcgatcctg gcgtcgatcc agctgttcgt gttcggcacc 540
tggctgccgc accgcccccg ccacgacgcg ttcccgacc gccataatgc gcggtcgtcg 600
cggatcagcg accccgtgtc gctgctgacc tgctttcact ttggtggtta tcatcacgaa 660
caccacctgc acccgacggt gccttgggtg cgcttcccc gcacccgcac caagggggac 720
accgcatga 729
```

```
<210> 6
<211> 242
<212> PRT
<213> Paracoccus haeundaesis
```

```
<220>
<223> crtW amino acid
```

```
<400> 6
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
1 5 10 15
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
Page 6
```

SubstCRFsequencelisting_2009Jul30.txt

20 25 30
 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
 35 40 45
 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala
 50 55 60
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn
 65 70 75 80
 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95
 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr
 100 105 110
 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
 115 120 125
 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
 130 135 140
 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
 145 150 155 160
 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
 165 170 175
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
 180 185 190
 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
 195 200 205
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
 210 215 220
 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
 225 230 235 240
 Thr Ala

<210> 7
 <211> 489
 <212> DNA
 <213> Paracoccus haeundaesis

<220>
 <223> crtZ gene

<400> 7
 atgaccaatt tcctgatcgt cgctgccacc gtgctggtga tggagttgac ggcctattcc 60
 gtccaccggtt ggatcatgca cggccccctg ggctggggct ggcacaagtc ccaccacgag 120
 gaacacgacc acgcgctgga aaagaacgac ctgtacggcc tggctctttgc ggtgatcgcc 180
 acggtgctgt tcacggtggg ctggatctgg gcgccggtcc tgtggtggat cgctttgggc 240
 atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctgggttca tcagcgctgg 300
 Page 7

SubstCRFsequencelisting_2009Jul30.txt

ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgctg 360
caccacgcgg tcgaggggacg cgaccattgc gtcagcttcg gcttcatcta tgcgccgccg 420
gtcgacaagc tgaagcagga cctgaagacg tcgggcgtgc tgcggggccga ggcgcaggag 480
cgcacgtga 489

<210> 8
<211> 162
<212> PRT
<213> Paracoccus haeundaesis

<220>
<223> crtZ amino acid

<400> 8
Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu
1 5 10 15
Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30
Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
35 40 45
Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe
50 55 60
Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly
65 70 75 80
Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val
85 90 95
His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
100 105 110
Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp
115 120 125
His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
130 135 140
Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu
145 150 155 160
Arg Thr

<210> 9
<211> 1161
<212> DNA
<213> Paracoccus haeundaesis

<220>
<223> crtY gene

<400> 9
gtgacccatg acgtgctgct ggcagggggcg ggccttgcca acgggctgat cgccctggcg 60
Page 8

SubstCRFsequencelisting_2009Jul30.txt

```

ctgcgcgcgg cgcgccccga cctgcgggtg ctgctgctgg atcatgcggc gggaccgtca 120
gacggccata cctgggtcctg ccacgacccc gatctgtcgc cgcactggct ggcgcggctg 180
aagccctgc gccgcgcca ctggcccgac caggaggtgc gctttccccg ccatgcccgg 240
cggctggcca ccggttacgg gtcgctggac ggggcggcgc tggcggatgc ggtggcccgg 300
tcgggcgccc agatccgctg gaacagcgac atcgccctgc tggatgaaca gggggcgacg 360
ctgtcctgcg gcacccggat cgaggcgggc gcggtcctgg acgggcgcgg cgcgcagccg 420
tcgcggcatc tgaccgtggg tttccagaaa ttcgtgggcg tcgagatcga gaccgactgc 480
ccccacggcg tgccccgccc gatgatcatg gacgcgaccg tcaccagca ggacgggtac 540
cgattcatct atctgtgcc cttctctccg acgcgcatcc tgatcgagga cactcgctat 600
tccgatggcg gcaatctgga cgacgacgcg ctggcggcgg cgtccacga ctatgcccg 660
cagcagggtt ggaccggggc cgaggtccgg cgcgaacgcg gcatcctgcc cattgcgctg 720
gcccatgacg cggcgggctt ctgggcccgat cagcggagg ggcctgttcc cgtgggactg 780
cgcgcggggg tctttcaccg ggtcaccggc tattcgtgc cctatgcggc gcagggtggc 840
gacgtggtgg cgggcctgtc cgggcccggc ggcaccgacg cgctgcgcgg cgccatccgc 900
gattacgcga tcgaccgggc acgccgtgac cgctttctgc gcctgctgaa ccggatgctg 960
ttccgcggct gcgcgcccga ccggcgctat accctgctgc agcggttcta ccgcatgccg 1020
catggactga tcgaacggtt ctatgccggc cggctgagcg tggcggatca gctgcgcac 1080
gtgaccggca agcctcccat tccccttggc acggccatcc gctgcctgcc cgaacgtccc 1140
ctgctgaagg aaaacgcatg a 1161

```

```

<210> 10
<211> 386
<212> PRT
<213> Paracoccus haeundaesis

```

```

<220>
<223> crtY amino acid

```

```

<400> 10
Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu
 1          5          10          15
Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu
          20          25          30
Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His
          35          40          45
Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg
          50          55          60
Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg
          65          70          75          80

```

SubstCRFsequencelisting_2009Jul30.txt

Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp
85 90 95

Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala
100 105 110

Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu
115 120 125

Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu
130 135 140

Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys
145 150 155 160

Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln
165 170 175

Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg
180 185 190

Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp
195 200 205

Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp
210 215 220

Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu
225 230 235 240

Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val
245 250 255

Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser
260 265 270

Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly
275 280 285

Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile
290 295 300

Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu
305 310 315 320

Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe
325 330 335

Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu
340 345 350

Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro
355 360 365

Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu
370 375 380

Asn Ala
385

SubstCRFsequencelisting_2009Jul30.txt

<211> 1506

<212> DNA

<213> Paracoccus haeundaesis

<220>

<223> crtI gene

<400> 11

```

atgaacgccc attcgcccgcc ggccaagacc gccatcgtga tcggcgcagg ctttggcggg    60
ctggccctgg ccacccgccc gcagtcgcgc ggcacgcgca ccaccctggc cgaggcccg    120
gacaagcccc gcggggcgcg ctagtctctg cacgatcagg gccatgtctt cgacgcgggc    180
ccgaccgtca tcaccgaccc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac    240
atggcgcgcg acgtgacgct gatgccggtg tcgcccttct atcgactgat gtggccgggc    300
gggaaggtct tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc    360
aaccgagcgc acctggaagg ataccgccc ttccgtgatt acgcggagga ggtgtatcag    420
gagggctacg tcaagctggg caccgtgccc ttctcaagc tgggccagat gctcaaggcc    480
gcgcccgcgc tgatgaagct ggaggcctat aagtccgtcc atgccaaggc cgcgaccttc    540
atcaaggacc cctatctgcg gcaggcgttt tcgtatcaca cgctgctggg gggcgggaat    600
cccttctcga ccagctcgat ctagcgctg atccacgcgc tggagcggcg cggcggggtc    660
tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg    720
cttggcggcc agatgatgct gaacgccaag gtcgcccgga tcgagaccga gggcgcgcg    780
accacgggcg tcaccctggc ggacgggccc tctttaaggg ccgacatggt cgccagcaac    840
ggcgacgtca tgcacaacta tcgcgacctg ctggggcaca cggcccgcgg gcagagccgc    900
gcgaaatcgc tggaccgcaa gcgctggtcc atgtcgttgt tcgtgctgca tttcggtctg    960
cgcgaggcgc ccaaggacat cgcgcatcac accatcctgt tcggcccccg ctacagggag   1020
ctggtcaacg agatcttcaa gggcccgaag ctggccgagg atttctcgct gtacctgcat   1080
tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg   1140
gccccgtgca cgcatctggg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat   1200
gccgaccgca tcctggcgct cctggaggag cggctgatcc cgaacctgcg cgccaacctg   1260
accacgacgc gcatcttcac gcccgccgat ttccgcagcg aactgaacgc ccatcacggc   1320
agcgcttctt cggtcgagcc gactctgacg caatccgcgt ggttccggcc gcacaaccgc   1380
gacaagacga tccgcaactt ctagctggtc ggcgcgggca cccatccggg cgcgggcatt   1440
ccgggcgctc tgggctcggc caaggccacg gccaggtga tgctgtccga cctggcgggc   1500
gcatga                                           1506

```

<210> 12

<211> 501

SubstCRFsequencelisting_2009Jul30.txt

<212> PRT

<213> Paracoccus haeundaesis

<220>

<223> crtI amino acid

<400> 12

```

Met Asn Ala His Ser Pro Ala Ala Lys Thr Ala Ile Val Ile Gly Ala
 1      5      10      15
Gly Phe Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ser Ala Gly Ile
 20      25      30
Ala Thr Thr Leu Val Glu Ala Arg Asp Lys Pro Gly Gly Arg Ala Tyr
 35      40      45
Val Trp His Asp Gln Gly His Val Phe Asp Ala Gly Pro Thr Val Ile
 50      55      60
Thr Asp Pro Asp Ala Leu Lys Glu Leu Trp Ala Leu Thr Gly Gln Asp
 65      70      75      80
Met Ala Arg Asp Val Thr Leu Met Pro Val Ser Pro Phe Tyr Arg Leu
 85      90      95
Met Trp Pro Gly Gly Lys Val Phe Asp Tyr Val Asn Glu Ala Asp Gln
100      105      110
Leu Glu Arg Gln Ile Ala Gln Phe Asn Pro Asp Asp Leu Glu Gly Tyr
115      120      125
Arg Arg Phe Arg Asp Tyr Ala Glu Glu Val Tyr Gln Glu Gly Tyr Val
130      135      140
Lys Leu Gly Thr Val Pro Phe Leu Lys Leu Gly Gln Met Leu Lys Ala
145      150      155      160
Ala Pro Ala Leu Met Lys Leu Glu Ala Tyr Lys Ser Val His Ala Lys
165      170      175
Val Ala Thr Phe Ile Lys Asp Pro Tyr Leu Arg Gln Ala Phe Ser Tyr
180      185      190
His Thr Leu Leu Val Gly Gly Asn Pro Phe Ser Thr Ser Ser Ile Tyr
195      200      205
Ala Leu Ile His Ala Leu Glu Arg Arg Gly Gly Val Trp Phe Ala Lys
210      215      220
Gly Gly Thr Asn Gln Leu Val Ala Gly Met Val Ala Leu Phe Glu Arg
225      230      235      240
Leu Gly Gly Gln Met Met Leu Asn Ala Lys Val Ala Arg Ile Glu Thr
245      250      255
Glu Gly Ala Arg Thr Thr Gly Val Thr Leu Ala Asp Gly Arg Ser Leu
260      265      270
Arg Ala Asp Met Val Ala Ser Asn Gly Asp Val Met His Asn Tyr Arg
275      280      285
Asp Leu Leu Gly His Thr Ala Arg Gly Gln Ser Arg Ala Lys Ser Leu
290      295      300

```

SubstCRFsequenceIisting_2009Jul30.txt

Asp Arg Lys Arg Trp Ser Met Ser Leu Phe Val Leu His Phe Gly Leu
 305 310 315 320
 Arg Glu Ala Pro Lys Asp Ile Ala His His Thr Ile Leu Phe Gly Pro
 325 330 335
 Arg Tyr Arg Glu Leu Val Asn Glu Ile Phe Lys Gly Pro Lys Leu Ala
 340 345 350
 Glu Asp Phe Ser Leu Tyr Leu His Ser Pro Cys Thr Thr Asp Pro Asp
 355 360 365
 Met Ala Pro Pro Gly Met Ser Thr His Tyr Val Leu Ala Pro Val Pro
 370 375 380
 His Leu Gly Arg Ala Glu Ile Asp Trp Ala Val Glu Gly Pro Arg Tyr
 385 390 395 400
 Ala Asp Arg Ile Leu Ala Ser Leu Glu Glu Arg Leu Ile Pro Asn Leu
 405 410 415
 Arg Ala Asn Leu Thr Thr Thr Arg Ile Phe Thr Pro Ala Asp Phe Ala
 420 425 430
 Ser Glu Leu Asn Ala His His Gly Ser Ala Phe Ser Val Glu Pro Ile
 435 440 445
 Leu Thr Gln Ser Ala Trp Phe Arg Pro His Asn Arg Asp Lys Thr Ile
 450 455 460
 Arg Asn Phe Tyr Leu Val Gly Ala Gly Thr His Pro Gly Ala Gly Ile
 465 470 475 480
 Pro Gly Val Val Gly Ser Ala Lys Ala Thr Ala Gln Val Met Leu Ser
 485 490 495
 Asp Leu Ala Gly Ala
 500

<210> 13
 <211> 915
 <212> DNA
 <213> Paracoccus haeundaesis

* <220>
 <223> crtB gene

<400> 13
 atgagcgatc tggctctgac ctcgaccgag gcgatcaccc aagggtcgca aagctttgcc 60
 acggcggcca agctgatgcc gccgggcatc cgcgacgaca cggtgatgct ctatgcctgg 120
 tgccgccacg cggatgacgt gatcgacggt caggccctgg gcagccgccc cgaggcggtg 180
 aacgaccgcg aggcgcggct ggacggcctg cgcgtcgaca cgctggcggc cctgcagggc 240
 gacggtccgg tgaccccgcc ctttgccgcg ctgcgcgcgg tggcgcggcg gcatgatttc 300
 ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac 360

SubstCRFsequencelisting_2009Jul30.txt

tatcgcacgc tggatgacgt gctggaatat tcctatcacg tcgcaggcat cgtcggcggtg 420
atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg 480
gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcatcggg 540
cggtgctatc tgccggggga ctggctggac caggcggggc cgcggatcga cgggccgggtg 600
ccgtcgccgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac 660
gcgtcggcgc ggggtgggtct ggcggatctg ccaccgcgt gcgcctggtc catcgccgcc 720
gcgctacgga tctatcgcg ccatcgggctg cgcattccga agagcgggcc gcaggcctat 780
cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggcgt cgggggctgg 840
gatgtcgcgc gatcacgcct gccggggggc ggcgtgtcgc ggcagggcct ctggacccgg 900
ccgcatcacg tctag 915

<210> 14
<211> 304
<212> PRT
<213> Paracoccus haeundaesis

<220>
<223> crtB amino acid

<400> 14
Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser
1 5 10 15
Gln Ser Phe Ala Thr Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Asp
20 25 30
Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile
35 40 45
Asp Gly Gln Ala Leu Gly Ser Arg Pro Glu Ala Val Asn Asp Pro Gln
50 55 60
Ala Arg Leu Asp Gly Leu Arg Val Asp Thr Leu Ala Ala Leu Gln Gly
65 70 75 80
Asp Gly Pro Val Thr Pro Pro Phe Ala Ala Leu Arg Ala Val Ala Arg
85 90 95
Arg His Asp Phe Pro Gln Ala Trp Pro Met Asp Leu Ile Glu Gly Phe
100 105 110
Ala Met Asp Val Glu Ala Arg Asp Tyr Arg Thr Leu Asp Asp Val Leu
115 120 125
Glu Tyr Ser Tyr His Val Ala Gly Ile Val Gly Val Met Met Ala Arg
130 135 140
Val Met Gly Val Arg Asp Asp Pro Val Leu Asp Arg Ala Cys Asp Leu
145 150 155 160
Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp
165 170 175

SubstCRFsequencelisting_2009Jul30.txt

Ala Arg Ile Gly Arg Cys Tyr Leu Pro Gly Asp Trp Leu Asp Gln Ala
180 185 190
Gly Ala Arg Ile Asp Gly Pro Val Pro Ser Pro Glu Leu Tyr Thr Val
195 200 205
Ile Leu Arg Leu Leu Asp Glu Ala Glu Pro Tyr Tyr Ala Ser Ala Arg
210 215 220
Val Gly Leu Ala Asp Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala
225 230 235 240
Ala Leu Arg Ile Tyr Arg Ala Ile Gly Leu Arg Ile Arg Lys Ser Gly
245 250 255
Pro Gln Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile
260 265 270
Gly Leu Leu Gly Val Gly Gly Trp Asp Val Ala Arg Ser Arg Leu Pro
275 280 285
Gly Ala Gly Val Ser Arg Gln Gly Leu Trp Thr Arg Pro His His Val
290 295 300

<210> 15
<211> 882
<212> DNA
<213> Paracoccus haeundaesis

<220>
<223> crtE gene

<400> 15
atgagacgag acgtcaaccc gatccacgcc acccttctgc agaccagact tgaggagatc 60
gccagggat tcggtgccgt gtcgcagccg ctcggcgcgg ccatgagcca tggcgcgctg 120
tcgtcgggca ggcggttccg cggcatgctg atgctgcttg cggcagaggc ctcgggcggg 180
gtctgcgaca cgatcgtcga cgccgcctgc gcggtcgaga tggatcatgc cgcacgctg 240
atcttcgacg acctgccctg catggacgat gccgggctgc gccgcggccg gcccgcgacc 300
catgtggcgc atggcgaaag ccgtgccgtg ctgggcggca tcgccctgat caccgaggca 360
atggccctgc tggccgggtgc gcgcggcgcg tcgggcacgg tcggggcgca gctggtgcgg 420
atcctgtcgc ggtccctggg gccgcagggc ctgtgcgccg gccaggacct ggacctgcac 480
gcggccaaga acggcgcggg ggtcgaacag gaacaggacc tgaagaccgg cgtgctgttc 540
atcgccgggc tggagatgct ggccgtgatc aaggagtctg acgccgagga gcagaccag 600
atgatcgact ttggccgtca gctgggcccgc gtgttccagt cctatgacga cctgctggac 660
gtcgtgggcg accaggcggc gcttggcaag gataccggtc gcgatgccgc ggccccggc 720
ccgcggcgcg gccttctggc cgtgtcagac ctgcagaacg tgtcccgtca ttacgaggcc 780
agccgcgccc aactggacgc gatgctgcgc agcaagcgcc ttcaggctcc ggaaatcgcg 840

gccctgctgg aacgggttct gccctacgcc gcgcgcgcct ag

<210> 16
 <211> 293
 <212> PRT
 <213> Paracoccus haeundaesis

<220>
 <223> crtE amino acid

<400> 16
 Met Arg Arg Asp Val Asn Pro Ile His Ala Thr Leu Leu Gln Thr Arg
 1 5 10 15
 Leu Glu Glu Ile Ala Gln Gly Phe Gly Ala Val Ser Gln Pro Leu Gly
 20 25 30
 Ala Ala Met Ser His Gly Ala Leu Ser Ser Gly Arg Arg Phe Arg Gly
 35 40 45
 Met Leu Met Leu Leu Ala Ala Glu Ala Ser Gly Gly Val Cys Asp Thr
 50 55 60
 Ile Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
 65 70 75 80
 Ile Phe Asp Asp Leu Pro Cys Met Asp Asp Ala Gly Leu Arg Arg Gly
 85 90 95
 Arg Pro Ala Thr His Val Ala His Gly Glu Ser Arg Ala Val Leu Gly
 100 105 110
 Gly Ile Ala Leu Ile Thr Glu Ala Met Ala Leu Leu Ala Gly Ala Arg
 115 120 125
 Gly Ala Ser Gly Thr Val Arg Ala Gln Leu Val Arg Ile Leu Ser Arg
 130 135 140
 Ser Leu Gly Pro Gln Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His
 145 150 155 160
 Ala Ala Lys Asn Gly Ala Gly Val Glu Gln Glu Gln Asp Leu Lys Thr
 165 170 175
 Gly Val Leu Phe Ile Ala Gly Leu Glu Met Leu Ala Val Ile Lys Glu
 180 185 190
 Phe Asp Ala Glu Glu Gln Thr Gln Met Ile Asp Phe Gly Arg Gln Leu
 195 200 205
 Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Val Gly Asp
 210 215 220
 Gln Ala Ala Leu Gly Lys Asp Thr Gly Arg Asp Ala Ala Ala Pro Gly
 225 230 235 240
 Pro Arg Arg Gly Leu Leu Ala Val Ser Asp Leu Gln Asn Val Ser Arg
 245 250 255
 His Tyr Glu Ala Ser Arg Ala Gln Leu Asp Ala Met Leu Arg Ser Lys
 260 265 270

SubstCRFsequencelisting_2009Jul30.txt

Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro
275 280 285

Tyr Ala Ala Arg Ala
290

<210> 17
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> forward primer for crt gene

<400> 17
gttcacgac tggggcatc

19

<210> 18
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> reverse primer for crt gene

<400> 18
tccactgacc ttgttgaca aattgccg

28

<210> 19
<211> 162
<212> PRT
<213> Alcaligenes sp.

<220>
<223> Alcaligenes sp. CrtZ Amino Acid

<400> 19
Met Thr Gln Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu
1 5 10 15

Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
20 25 30

Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
35 40 45

Asn Asp Leu Tyr Gly Val Val Phe Ala Val Leu Ala Thr Ile Leu Phe
50 55 60

Thr Val Gly Ala Tyr Trp Trp Pro Val Leu Trp Trp Ile Ala Leu Gly
65 70 75 80

Met Thr Val Tyr Gly Leu Ile Tyr Phe Ile Leu His Asp Gly Leu Val
85 90 95

His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Arg Gly Tyr Phe Arg
100 105 110

SubstCRFsequencelisting_2009Jul30.txt

Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp
 115 120 125
 His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
 130 135 140
 Lys Gln Asp Leu Lys Arg Ser Gly Val Leu Arg Pro Gln Asp Glu Arg
 145 150 155 160

Pro Ser

<210> 20
 <211> 242
 <212> PRT
 <213> Alcaligenes sp.

<220>
 <223> Alcaligenes sp. Crtw Amino Acid

<400> 20
 Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu
 1 5 10 15
 Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe
 20 25 30
 Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu
 35 40 45
 Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
 50 55 60
 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn
 65 70 75 80
 Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp
 85 90 95
 Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr
 100 105 110
 Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
 115 120 125
 Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
 130 135 140
 Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
 145 150 155 160
 Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe
 165 170 175
 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro
 180 185 190
 Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu
 195 200 205
 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His

210

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly
225 230 235 240

Arg Ala

<210> 21
<211> 258
<212> PRT
<213> Bradyrhizobium sp.

<220>
<223> Bradyrhizobium sp. CrtW Amino Acid

<400> 21
Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg
1 5 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro
35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln
50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His
65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp
115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe
130 135 140

Phe Leu His Tyr Phe Gly Trp Lys Gln Val Ala Ile Ile Ala Ala Val
145 150 155 160

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu
165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp
195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Leu Ser Leu Leu
210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp
225 230 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
 245 250 255

Arg Asp

<210> 22
 <211> 295
 <212> PRT
 <213> Flavobacterium sp.

<220>
 <223> Flavobacterium sp. CrtE Amino Acid

<400> 22
 Met Thr Pro Lys Gln Gln Phe Pro Leu Arg Asp Leu Val Glu Ile Arg
 1 5 10 15
 Leu Ala Gln Ile Ser Gly Gln Phe Gly Val Val Ser Ala Pro Leu Gly
 20 25 30
 Ala Ala Met Ser Asp Ala Ala Leu Ser Pro Gly Lys Arg Phe Arg Ala
 35 40 45
 Val Leu Met Leu Met Val Ala Glu Ser Ser Gly Gly Val Cys Asp Ala
 50 55 60
 Met Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
 65 70 75 80
 Ile Phe Asp Asp Met Pro Cys Met Asp Asp Ala Arg Thr Arg Arg Gly
 85 90 95
 Gln Pro Ala Thr His Val Ala His Gly Glu Gly Arg Ala Val Leu Ala
 100 105 110
 Gly Ile Ala Leu Ile Thr Glu Ala Met Arg Ile Leu Gly Glu Ala Arg
 115 120 125
 Gly Ala Thr Pro Asp Gln Arg Ala Arg Leu Val Ala Ser Met Ser Arg
 130 135 140
 Ala Met Gly Pro Val Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His
 145 150 155 160
 Ala Pro Lys Asp Ala Ala Gly Ile Glu Arg Glu Gln Asp Leu Lys Thr
 165 170 175
 Gly Val Leu Phe Val Ala Gly Leu Glu Met Leu Ser Ile Ile Lys Gly
 180 185 190
 Leu Asp Lys Ala Glu Thr Glu Gln Leu Met Ala Phe Gly Arg Gln Leu
 195 200 205
 Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Ile Gly Asp
 210 215 220
 Lys Ala Ser Thr Gly Lys Asp Thr Ala Arg Asp Thr Ala Ala Pro Gly
 225 230 235 240
 Pro Lys Gly Gly Leu Met Ala Val Gly Gln Met Gly Asp Val Ala Gln
 Page 20

SubstCRFsequencelisting_2009Jul30.txt

245 250 255
 His Tyr Arg Ala Ser Arg Ala Gln Leu Asp Glu Leu Met Arg Thr Arg
 260 265 270
 Leu Phe Arg Gly Gly Gln Ile Ala Asp Leu Leu Ala Arg Val Leu Pro
 275 280 285
 His Asp Ile Arg Arg Ser Ala
 290 295

<210> 23
 <211> 303
 <212> PRT
 <213> Flavobacterium sp.

<220>
 <223> Flavobacterium sp. CrtB Amino Acid

<400> 23
 Met Thr Asp Leu Thr Ala Thr Ser Glu Ala Ala Ile Ala Gln Gly Ser
 1 5 10 15
 Gln Ser Phe Ala Gln Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Glu
 20 25 30
 Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile
 35 40 45
 Asp Gly Gln Val Met Gly Ser Ala Pro Glu Ala Gly Gly Asp Pro Gln
 50 55 60
 Ala Arg Leu Gly Ala Leu Arg Ala Asp Thr Leu Ala Ala Leu His Glu
 65 70 75 80
 Asp Gly Pro Met Ser Pro Pro Phe Ala Ala Leu Arg Gln Val Ala Arg
 85 90 95
 Arg His Asp Phe Pro Asp Leu Trp Pro Met Asp Leu Ile Glu Gly Phe
 100 105 110
 Ala Met Asp Val Ala Asp Arg Glu Tyr Arg Ser Leu Asp Asp Val Leu
 115 120 125
 Glu Tyr Ser Tyr His Val Ala Gly Val Val Gly Val Met Met Ala Arg
 130 135 140
 Val Met Gly Val Gln Asp Asp Ala Val Leu Asp Arg Ala Cys Asp Leu
 145 150 155 160
 Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp
 165 170 175
 Ala Ala Ile Gly Arg Cys Tyr Leu Pro Ala Asp Trp Leu Ala Glu Ala
 180 185 190
 Gly Ala Thr Val Glu Gly Pro Val Pro Ser Asp Ala Leu Tyr Ser Val
 195 200 205
 Ile Ile Arg Leu Leu Asp Ala Ala Glu Pro Tyr Tyr Ala Ser Ala Arg
 210 215 220

SubstCRFsequencelisting_2009Jul30.txt

Gln Gly Leu Pro His Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala
 225 230 235 240
 Ala Leu Arg Ile Tyr Arg Ala Ile Gly Thr Arg Ile Arg Gln Gly Gly
 245 250 255
 Pro Glu Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile
 260 265 270
 Gly Leu Leu Ala Arg Gly Gly Leu Asp Ala Ala Ala Ser Arg Leu Arg
 275 280 285
 Gly Gly Glu Ile Ser Arg Asp Gly Leu Trp Thr Arg Pro Arg Ala
 290 295 300

<210> 24
 <211> 494
 <212> PRT
 <213> Flavobacterium sp.

<220>
 <223> Flavobacterium sp. CrtI Amino Acid

<400> 24
 Met Ser Ser Ala Ile Val Ile Gly Ala Gly Phe Gly Gly Leu Ala Leu
 1 5 10 15
 Ala Ile Arg Leu Gln Ser Ala Gly Ile Ala Thr Thr Ile Val Glu Ala
 20 25 30
 Arg Asp Lys Pro Gly Gly Arg Ala Tyr Val Trp Asn Asp Gln Gly His
 35 40 45
 Val Phe Asp Ala Gly Pro Thr Val Val Thr Asp Pro Asp Ser Leu Arg
 50 55 60
 Glu Leu Trp Ala Leu Ser Gly Gln Pro Met Glu Arg Asp Val Thr Leu
 65 70 75 80
 Leu Pro Val Ser Pro Phe Tyr Arg Leu Thr Trp Ala Asp Gly Arg Ser
 85 90 95
 Phe Glu Tyr Val Asn Asp Asp Asp Glu Leu Ile Arg Gln Val Ala Ser
 100 105 110
 Phe Asn Pro Ala Asp Val Asp Gly Tyr Arg Arg Phe His Asp Tyr Ala
 115 120 125
 Glu Glu Val Tyr Arg Glu Gly Tyr Leu Lys Leu Gly Thr Thr Pro Phe
 130 135 140
 Leu Lys Leu Gly Gln Met Leu Asn Ala Ala Pro Ala Leu Met Arg Leu
 145 150 155 160
 Gln Ala Tyr Arg Ser Val His Ser Met Val Ala Arg Phe Ile Gln Asp
 165 170 175
 Pro His Leu Arg Gln Ala Phe Ser Phe His Thr Leu Leu Val Gly Gly
 180 185 190

SubstCRFsequencelisting_2009Jul30.txt

Asn Pro Phe Ser Thr Ser Ser Ile Tyr Ala Leu Ile His Ala Leu Glu
195 200 205

Arg Arg Gly Gly Val Trp Phe Ala Lys Gly Gly Thr Asn Gln Leu Val
210 215 220

Ala Gly Met Val Ala Leu Phe Glu Arg Leu Gly Gly Thr Leu Leu Leu
225 230 235 240

Asn Ala Arg Val Thr Arg Ile Asp Thr Glu Gly Asp Arg Ala Thr Gly
245 250 255

Val Thr Leu Leu Asp Gly Arg Gln Leu Arg Ala Asp Thr Val Ala Ser
260 265 270

Asn Gly Asp Val Met His Ser Tyr Arg Asp Leu Leu Gly His Thr Arg
275 280 285

Arg Gly Arg Thr Lys Ala Ala Ile Leu Asn Arg Gln Arg Trp Ser Met
290 295 300

Ser Leu Phe Val Leu His Phe Gly Leu Ser Lys Arg Pro Glu Asn Leu
305 310 315 320

Ala His His Ser Val Ile Phe Gly Pro Arg Tyr Lys Gly Leu Val Asn
325 330 335

Glu Ile Phe Asn Gly Pro Arg Leu Pro Asp Asp Phe Ser Met Tyr Leu
340 345 350

His Ser Pro Cys Val Thr Asp Pro Ser Leu Ala Pro Glu Gly Met Ser
355 360 365

Thr His Tyr Val Leu Ala Pro Val Pro His Leu Gly Arg Ala Asp Val
370 375 380

Asp Trp Glu Ala Glu Ala Pro Gly Tyr Ala Glu Arg Ile Phe Glu Glu
385 390 395 400

Leu Glu Arg Arg Ala Ile Pro Asp Leu Arg Lys His Leu Thr Val Ser
405 410 415

Arg Ile Phe Ser Pro Ala Asp Phe Ser Thr Glu Leu Ser Ala His His
420 425 430

Gly Ser Ala Phe Ser Val Glu Pro Ile Leu Thr Gln Ser Ala Trp Phe
435 440 445

Arg Pro His Asn Arg Asp Arg Ala Ile Pro Asn Phe Tyr Ile Val Gly
450 455 460

Ala Gly Thr His Pro Gly Ala Gly Ile Pro Gly Val Val Gly Ser Ala
465 470 475 480

Lys Ala Thr Ala Gln Val Met Leu Ser Asp Leu Ala Val Ala
485 490

<210> 25
<211> 382
<212> PRT
<213> Flavobacterium sp.

SubstCRFsequencelisting_2009Jul30.txt

<220>

<223> Flavobacterium sp. CrtY Amino Acid

<400>

```

      25
Met Ser His Asp Leu Leu Ile Ala Gly Ala Gly Leu Ser Gly Ala Leu
  1      5      10      15
Ile Ala Leu Ala Val Arg Asp Arg Arg Pro Asp Ala Arg Ile Val Met
      20      25      30
Leu Asp Ala Arg Ser Gly Pro Ser Asp Gln His Thr Trp Ser Cys His
      35      40      45
Asp Thr Asp Leu Ser Pro Glu Trp Leu Ala Arg Leu Ser Pro Ile Arg
      50      55      60
Arg Gly Glu Trp Thr Asp Gln Glu Val Ala Phe Pro Asp His Ser Arg
      65      70      75      80
Arg Leu Thr Thr Gly Tyr Gly Ser Ile Glu Ala Gly Ala Leu Ile Gly
      85      90      95
Leu Leu Gln Gly Val Asp Leu Arg Trp Asn Thr His Val Ala Thr Leu
      100      105      110
Asp Asp Thr Gly Ala Thr Leu Thr Asp Gly Ser Arg Ile Glu Ala Ala
      115      120      125
Cys Val Ile Asp Ala Arg Gly Ala Val Glu Thr Pro His Leu Thr Val
      130      135      140
Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Ala Pro His
      145      150      155      160
Gly Val Glu Arg Pro Met Ile Met Asp Ala Thr Val Pro Gln Met Asp
      165      170      175
Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg Ile Leu
      180      185      190
Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asp Leu Asp Asp Gly Ala
      195      200      205
Leu Ala Gln Ala Ser Leu Asp Tyr Ala Ala Arg Arg Gly Trp Thr Gly
      210      215      220
Gln Glu Met Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu Ala His
      225      230      235      240
Asp Ala Ile Gly Phe Trp Arg Asp His Ala Gln Gly Ala Val Pro Val
      245      250      255
Gly Leu Gly Ala Gly Leu Phe His Pro Val Thr Gly Tyr Ser Leu Pro
      260      265      270
Tyr Ala Ala Gln Val Ala Asp Ala Ile Ala Ala Arg Asp Leu Thr Thr
      275      280      285
Ala Ser Ala Arg Arg Ala Val Arg Gly Trp Ala Ile Asp Arg Ala Asp
      290      295      300
Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu Phe Arg Gly Cys

```

SubstCRFsequencelisting_2009jul30.txt

305 310 315 320
Pro Pro Asp Arg Arg Tyr Arg Leu Leu Gln Arg Phe Tyr Arg Leu Pro
325 330 335
Gln Pro Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu Thr Leu Ala Asp
340 345 350
Arg Leu Arg Ile Val Thr Gly Arg Pro Pro Ile Pro Leu Ser Gln Ala
355 360 365
Val Arg Cys Leu Pro Glu Arg Pro Leu Leu Gln Glu Arg Ala
370 375 380